

Off-the-grid Two-Dimensional Line Spectral Estimation With Prior Information

Iman Valiulahi, Hamid Fathi, Sajad Daei, and Farzan Haddadi

Abstract—In this paper, we provide a method to recover off-the-grid frequencies of a signal in two-dimensional (2-D) line spectral estimation. Most of the literature in this field focuses on the case in which the only information is spectral sparsity in a continuous domain and does not consider prior information. However, in many applications such as radar and sonar, one has extra information about the spectrum of the signal of interest. The common way of accommodating prior information is to use weighted atomic norm minimization. We present a new semidefinite program using the theory of positive trigonometric polynomials that incorporate this prior information into 2-D line spectral estimation. Specifically, we assume prior knowledge of 2-D frequency subbands in which signal frequency components are located. Our approach improves the recovery performance compared with the previous work that does not consider prior information. Through numerical experiments, we find out that the amount of this improvement depends on prior information we have about the locations of the true frequencies.

Index Terms—Continuous compressed sensing, Prior information, atomic norm, Line spectral estimation.

I. INTRODUCTION

In recent years, there is a growing interest in continuous compressed sensing (CS) as it appears in many signal processing applications such as radar, sonar, array processing, seismology, and remote sensing [1], [2]. The problem is to recover a signal which is known to be sparse in a continuous domain. In this paper, we investigate line spectral estimation as a special case of continuous CS where the signal of interest is sparse in the continuous frequency domain. Conventional approaches in this field are based on super-resolution methods such as 2-D MUSIC [3], and Matrix Enhancement Matrix Pencil (MEMP) [2]. However, these approaches are sensitive to noise and outliers. Moreover, the number of signal frequency components are assumed to be known a priori, while they are not known in practice.

In compressed sensing, the goal is to find a sparse vector under the condition that one has some incomplete random linear observations known as l_0 problem [4], [5]. Since the l_0 problem is NP-hard one can use the best convex relaxation form of l_0 problem, which is the l_1 norm minimization problem. The required number of measurements for the l_1 problem to succeed is $\mathcal{O}(r \log n)$. Unlike the previous approaches, l_1 problem does not need the number of non-zero entries of the signal. In continuation of the previous work, Baraniuk in [6] investigated the reconstruction of sparse signals in the discrete fourier transform (DFT) dictionary. In such cases, the classical l_1 problem suffers from basis mismatch as it considers the signal is sparse in the discrete domain while in

many applications, the signal lies in a continuous dictionary [7], [8].

Tang et al. in [9] presents a new concept of “Compressed sensing off-the-grid”. They assumed the frequencies to take arbitrary values in the signal frequency domain and show that the number of necessary measurements is $\mathcal{O}(r \log r \log n)$ for perfect recovery.

2-D line spectral estimation based on convex optimization methods fall into two categories: atomic norm minimization methods introduced first by Chi et al. [10] that is an extended version of [9], and matrix completion methods known as Enhanced Matrix Completion (EMaC) [11].

In the previous works in off-the-grid 2-D line spectral estimation, the only available information is spectral sparsity. However, in practice extra information is often available e.g., in radar and sonar applications, targets are often change their locations very slowly. Therefore, prior approximate knowledge of their locations are available. This can translate to a prior inexact information about the locations of sparse coefficients in the fourier domain when we use array processing techniques to localize the targets. A major question in 2-D line spectral estimation is whether this additional information can improve the recovery performance. In this paper, the prior information is in the form of some 2-D frequency intervals that are expected to contain the true signal frequency components. We exploit this prior knowledge using weighted atomic norm optimization problem. Dual of weighted atomic norm is converted to a semidefinite program using the well-known theory of positive trigonometric polynomials. We experimentally find out that the more precise is the prior information, the better we can estimate the signal frequencies.

II. SYSTEM MODEL

Consider a data matrix $\mathbf{X} \in \mathbb{C}^{n_1 \times n_2}$. Without loss of generality assume that $n_1 = n_2 = n$. Each element of the matrix \mathbf{X} can be written as a mixture of r exponentials as follows:

$$x_{\mathbf{k}} = \frac{1}{n} \sum_{i=1}^r d_i e^{j2\pi \mathbf{f}_i^T \mathbf{k}}, \quad (1)$$

where $d_i = |d_i|e^{j\phi_i}$ is an arbitrary complex amplitude of each exponential with $\phi_i \in [0, 2\pi)$, $J = \{0, \dots, n-1\} \times \{0, \dots, n-1\}$ represents all of indices of the signal, $\mathbf{k} = (k_1, k_2) \in J$ and

each frequency component $\mathbf{f} = (f_{1i}, f_{2i}) \in [0, 1] \times [0, 1]$, $i = 1, \dots, r$. The frequency atoms are defined as:

$$\begin{aligned} \mathbf{a}(f_{1i}) &:= \frac{1}{\sqrt{n}} [1, \dots, y_i^{n-1}]^T, \\ \mathbf{a}(f_{2i}) &:= \frac{1}{\sqrt{n}} [1, \dots, z_i^{n-1}]^T, \end{aligned} \quad (2)$$

where $y_i = e^{j2\pi f_{1i}}$ and $z_i = e^{j2\pi f_{2i}}$. The matrix \mathbf{X} in (1) can be written in the following form:

$$\mathbf{X} = \mathbf{Y} \mathbf{D} \mathbf{Z}^T, \quad (3)$$

where:

$$\mathbf{Y} := [\mathbf{a}(f_{11}), \dots, \mathbf{a}(f_{1r})] \in \mathbb{C}^{n \times r}, \quad (4)$$

$$\mathbf{Z} := [\mathbf{a}(f_{21}), \dots, \mathbf{a}(f_{2r})] \in \mathbb{C}^{n \times r}, \quad (5)$$

and

$$\mathbf{D} := \text{diag}([d_1, \dots, d_r]) = \text{diag}(\mathbf{d}) \in \mathbb{C}^{r \times r}. \quad (6)$$

Notice that $\mathbf{x} = \text{vec}(\mathbf{X}^T) \in \mathbb{C}^{n^2}$ is the vectorized form of \mathbf{X} that can be written as follows:

$$\begin{aligned} \mathbf{x} &= (\mathbf{Y} \otimes \mathbf{Z}) \mathbf{d} = \sum_{i=1}^r d_i \mathbf{a}(f_{1i}) \otimes \mathbf{a}(f_{2i}) \\ &= \sum_{i=1}^r d_i \mathbf{c}(f_i) = \sum_{i=1}^r |d_i| \mathbf{c}(f_i, \phi_i), \end{aligned} \quad (7)$$

where \otimes represents the kronecker product and $\mathbf{c}(f_i, \phi_i) = \mathbf{c}(f_{1i}, f_{2i}) e^{j\phi_i} \triangleq \mathbf{a}(f_{1i}) \otimes \mathbf{a}(f_{2i}) e^{j\phi_i} \in \mathbb{C}^{n^2}$. By the above definition it is easy to see that $\|\mathbf{c}(f_i)\|_2 = 1$.

In many linear inverse problems, the goal is to minimize a structure inducing norm with the constraint of observations [12]. The unit ball of the structure inducing norm is the convex hull of a special atomic set. For instance, convex hull of the standard basis vectors and their outer product as rank-one matrices form the unit ball of the l_1 norm and nuclear norm, respectively. We will use the following special atomic norm defined as:

$$\|\mathbf{x}\|_{\mathcal{A}} := \inf_{\mathbf{f}_i \in [0,1] \times [0,1], d_i \in \mathbb{C}} \left\{ \sum_i |d_i| \mid \mathbf{x} = \sum_i d_i \mathbf{c}_i(\mathbf{f}_i) \right\}, \quad (8)$$

where the atomic set \mathcal{A} is given by:

$$\mathcal{A} := \{\mathbf{c}(\mathbf{f}) \mid \mathbf{f} \in [0, 1] \times [0, 1]\}.$$

A common way for recovering a sparse signal with partial observations is the atomic norm optimization problem which is defined as:

$$\min_{\mathbf{x}^* \in \mathbb{C}^{n^2}} \|\mathbf{x}^*\|_{\mathcal{A}} \quad \text{subject to} \quad \mathcal{P}_T(\mathbf{x}^*) = \mathcal{P}_T(\mathbf{x}), \quad (9)$$

where T is the set of observed entries, and $\mathcal{P}_T(\mathbf{x})$ is a uniform sampling operator of $\mathbf{x} \in \mathbb{C}^{n^2}$.

It is often convenient to recover \mathbf{x} indirectly by first finding the true frequencies using the dual problem of (9) and then

the problem reduces to a linear system of equations that can be easily solved. The dual problem of (9) is as follows:

$$\begin{aligned} &\max_{\mathbf{q} \in \mathbb{C}^{n^2}, \mathbf{Q} \in \mathbb{C}^{n^2 \times n^2}} \langle \mathbf{x}, \mathbf{q} \rangle_{\mathbb{R}} \\ &\text{subject to} \quad \delta_{\mathbf{k}} = \text{tr}[\Theta_{\mathbf{k}} \mathbf{Q}], \quad \mathbf{k} \in \mathcal{H} \end{aligned} \quad (10)$$

$$\begin{bmatrix} \mathbf{Q} & \mathbf{q} \\ \mathbf{q}^* & \mathbf{1} \end{bmatrix} \succeq 0, \quad \mathbf{q}_{T^c} = 0,$$

where T^c is the complement of the set T , \mathbf{q} is the dual variable, $\langle \cdot, \cdot \rangle_{\mathbb{R}}$ denotes the real part of the inner product, \mathcal{H} is a half-space, $\Theta_{\mathbf{k}}$ is an elementary Toeplitz matrix with unit values on its \mathbf{k} th diagonal and zeros elsewhere, $\Theta_{\mathbf{k}} = \Theta_{k_2} \otimes \Theta_{k_1}$, $\delta_{\mathbf{k}} = 1$ if $\mathbf{k} = (0, 0)$, otherwise $\delta_{\mathbf{k}} = 0$, and the semidefinite constraint in (10) imposes $\mathbf{Q} \succeq 0$. Since the Slater's condition is satisfied in the primal problem (9), strong duality holds between (10) and (9).

Regarding the fact that the signal frequency components are the peaks of the following dual polynomial with unit modulus, one can estimate the true frequencies [9], [13]:

$$Q(\mathbf{f}) = \sum_{\mathbf{k} \in \mathcal{H}} q_{\mathbf{k}} e^{-j2\pi \mathbf{f}^T \mathbf{k}}. \quad (11)$$

In many applications, there exists some prior information about the parameters of the signal. Specifically, in this paper, the prior information is assumed in the signal spectrum. More precisely, assume that there exist p subbands in the frequency domain of the signal which with probabilities $\{P_F(\mathbf{f}_{P_i})\}_{i=1}^p$ are expected to have the true frequency components (Fig.1). Our approach is to use weighted atomic norm problem to exploit prior knowledge in this problem which is defined as follows:

$$\begin{aligned} &\min_{\mathbf{x}^*} \|\mathbf{x}^*\|_{w\mathcal{A}} \\ &\text{subject to} \quad \mathcal{P}_T(\mathbf{x}^*) = \mathcal{P}_T(\mathbf{x}), \end{aligned} \quad (12)$$

with $\|\mathbf{x}\|_{w\mathcal{A}}$ defined as:

$$\|\mathbf{x}\|_{w\mathcal{A}} := \inf_{\mathbf{f}_i \in [0,1] \times [0,1], d_i \in \mathbb{C}} \left\{ \sum_i w_i |d_i| \mid \mathbf{x} = \sum_i d_i \mathbf{c}_i(\mathbf{f}_i) \right\}, \quad (13)$$

where $\mathbf{w} = \{w_1, \dots, w_s\}$ is the vector of weights. Now we define $w(\mathbf{f}) = w_1 = \dots = w_s$ as a 2-D piecewise constant function associated with \mathbf{w} in each subband. The value of $w(\mathbf{f}_{P_i})$ in subband P_i is $\frac{1}{P_F(\mathbf{f}_{P_i})}$.

Although, in contrast to weighted l_1 minimization, there does not exist SDP formulation as in [9, Equation 2.6] for (12) [14], however, the dual problem of weighted atomic norm minimization can be used to recover the signal frequency components. The dual problem associated with (12) is given by:

$$\begin{aligned} &\max_{\mathbf{q}} \langle \mathbf{q}, \mathbf{x} \rangle_{\mathbb{R}} \\ &\text{subject to} \quad \|\mathbf{q}\|_{w\mathcal{A}}^* \leq 1, \\ &\mathbf{q}_{T^c} = 0, \end{aligned} \quad (14)$$

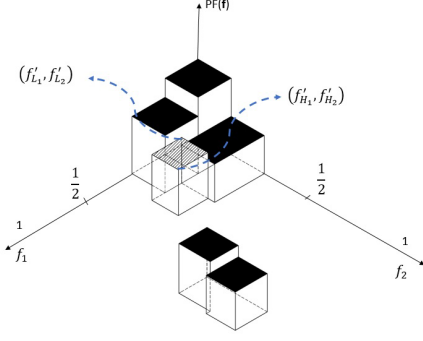


Fig. 1. Prior distribution of frequency components. The signal spectrum has been normalized to $[0, 1) \times [0, 1)$

where $\|\mathbf{q}\|_{w\mathcal{A}}^*$ is the dual norm of $\|\mathbf{q}\|_{w\mathcal{A}}$ which is defined as:

$$\begin{aligned} \|\mathbf{q}\|_{w\mathcal{A}}^* &= \sup_{\|\hat{\mathbf{x}}\|_{w\mathcal{A}} \leq 1} \langle \mathbf{q}, \hat{\mathbf{x}} \rangle_{\mathbb{R}} \\ &= \sup_{\phi \in [0, 2\pi], \mathbf{f} \in [0, 1) \times [0, 1)} \left\langle \mathbf{q}, e^{j\phi} \mathbf{c}(\mathbf{f}) \frac{1}{w(\mathbf{f})} \right\rangle_{\mathbb{R}} \\ &= \sup_{\mathbf{f} \in [0, 1) \times [0, 1)} \left| \left\langle \mathbf{q}, \mathbf{c}(\mathbf{f}) \frac{1}{w(\mathbf{f})} \right\rangle \right|. \end{aligned} \quad (15)$$

Replacing (15) in (14) gives:

$$\begin{aligned} &\max_{\mathbf{q}} \langle \mathbf{q}, \mathbf{x} \rangle_{\mathbb{R}} \\ &\text{subject to } \sup_{\mathbf{f} \in [0, 1) \times [0, 1)} |\langle \mathbf{q}, \mathbf{c}(\mathbf{f}) \frac{1}{w(\mathbf{f})} \rangle| \leq 1, \\ &\mathbf{q}_{T^c} = 0. \end{aligned} \quad (16)$$

Without any prior knowledge $w(\mathbf{f}) = 1 : \forall \mathbf{f} \in [0, 1) \times [0, 1)$, problem (10) is a special case of the problem (16).

Regarding our prior information, problem (16) converts to the following problem with p inequality constraints:

$$\begin{aligned} &\max_{\mathbf{q}} \langle \mathbf{q}, \mathbf{x} \rangle_{\mathbb{R}} \\ &\text{subject to } \sup_{\mathbf{f} \in P_i} |\langle \mathbf{q}, \mathbf{c}(\mathbf{f}) \rangle| \leq w(\mathbf{f}_{P_i}), \quad i = 1, \dots, p, \\ &\mathbf{q}_{T^c} = 0. \end{aligned} \quad (17)$$

The above relation generalizes the dual problem of 1-D weighted atomic norm in [15] to 2-D.

The following conditions and the positive trigonometric polynomial theory [16] are necessary to convert the inequality constraints (17) to linear matrix inequalities. Let $[f'_{L1}, f'_{H1}] \times [f'_{L2}, f'_{H2}]$ a single frequency band (the hatched subband in Fig. 1).

$$\begin{cases} f_{L_i} = f'_{L_i} & : 0 \leq f'_{L_i} \leq 0.5 \\ f_{H_i} = f'_{H_i} & : 0 \leq f'_{H_i} \leq 0.5 \\ f_{L_i} = 1 - f'_{H_i} & : 0.5 < f'_{H_i} < 1 \\ f_{H_i} = 1 - f'_{L_i} & : 0.5 < f'_{L_i} < 1 \\ f_{L_i} = 1 - f'_{H_i} & : 0 \leq f'_{L_i} \leq 0.5 \\ f_{H_i} = 0.5 & : 0.5 < f'_{H_i} \leq 1 \end{cases} \quad (18)$$

$$\begin{aligned} &\cos \omega_i - \cos \omega_{H_i} \geq 0, \\ &-\cos \omega_i + \cos \omega_{L_i} \geq 0, \end{aligned} \quad (19)$$

where $\omega_{H_i} = 2\pi f_{H_i}$, $\omega_{L_i} = 2\pi f_{L_i}$ and $i \in \{1, 2\}$. Notice that the complement of the domain $[f'_{L1}, f'_{H1}] \times [f'_{L2}, f'_{H2}]$ is obtained by changing the sign of the inequalities in (19).

It is well-known in the positive trigonometric polynomial theory [16] that if:

$$\begin{aligned} &\sup |\langle \mathbf{q}, \mathbf{c}(\mathbf{f}) \rangle| \leq \gamma \\ &\forall \mathbf{f} \in [f'_{L1}, f'_{H1}] \times [f'_{L2}, f'_{H2}], \end{aligned} \quad (20)$$

then there exist matrices $G_l \succeq 0, l \in \{1, \dots, 4\}$, such that:

$$\begin{aligned} \mathcal{G}_{\mathbf{k}, [f'_{L1}, f'_{H1}] \times [f'_{L2}, f'_{H2}]} &\triangleq \text{tr}[(\Theta_{k_2} \otimes \Theta_{k_1}) \mathbf{G}_0] + \mathcal{T}_{\mathbf{k}, 0.5, -\cos \omega_{H_1}}(\mathbf{G}_1) \\ &+ \mathcal{T}_{\mathbf{k}, -0.5, \cos \omega_{L_1}}(\mathbf{G}_2) + \mathcal{L}_{\mathbf{k}, 0.5, -\cos \omega_{H_2}}(\mathbf{G}_3) \\ &+ \mathcal{L}_{\mathbf{k}, -0.5, \cos \omega_{L_2}}(\mathbf{G}_4) = \delta_{\mathbf{k}}, \\ &\mathbf{k} \in \mathcal{H}, \quad \begin{bmatrix} \mathbf{G}_0 & \frac{1}{\gamma} \mathbf{q} \\ \frac{1}{\gamma} \mathbf{q}^* & 1 \end{bmatrix} \succeq 0, \end{aligned} \quad (21)$$

where $\mathbf{G}_0 \in \mathbb{C}^{n^2 \times n^2}$ and $\{\mathbf{G}_i\}_{i=1}^4 \in \mathbb{C}^{(n-1)^2 \times (n-1)^2}$. Moreover, $\mathcal{T}_{\mathbf{k}, d_0, d_1}(\mathbf{G}_1)$ and $\mathcal{L}_{\mathbf{k}, r_0, r_1}(\mathbf{G}_2)$ are defined as:

$$\begin{aligned} \mathcal{T}_{\mathbf{k}, d_0, d_1}(\mathbf{G}_1) &\triangleq \text{tr}[(d_1 \Theta_{\mathbf{k}_2} \otimes \Theta_{\mathbf{k}_1-1} + d_0 \Theta_{\mathbf{k}_2} \otimes \Theta_{\mathbf{k}_1} \\ &+ d_1 \Theta_{\mathbf{k}_2} \otimes \Theta_{\mathbf{k}_1+1}) \mathbf{G}_1], \end{aligned} \quad (22)$$

$$\begin{aligned} \mathcal{L}_{\mathbf{k}, r_0, r_1}(\mathbf{G}_2) &\triangleq \text{tr}[(r_1 \Theta_{\mathbf{k}_2-1} \otimes \Theta_{\mathbf{k}_1} + r_0 \Theta_{\mathbf{k}_2} \otimes \Theta_{\mathbf{k}_1} \\ &+ r_1 \Theta_{\mathbf{k}_2+1} \otimes \Theta_{\mathbf{k}_1}) \mathbf{G}_2]. \end{aligned} \quad (23)$$

The inequality in (21) implies that $\mathbf{G}_0 \succeq 0$. After applying (21) to each inequality constraint in (17), the problem (17) changes to the following problem:

$$\begin{aligned} &\max_{\mathbf{q}, \mathbf{G}_{01}, \dots, \mathbf{G}_{4p}} \langle \mathbf{q}, \mathbf{x} \rangle_{\mathbb{R}} \\ &\text{subject to } \mathcal{G}_{k_i, \mathbf{f}} = \delta_{\mathbf{k}_i}, \quad \mathbf{f} \in p_i, \\ &\mathbf{k}_i \in \{1 - n, \dots, n - 1\} \times \{1 - n, \dots, n - 1\}, \\ &\begin{bmatrix} \mathbf{G}_{0i} & \frac{1}{w_i(\mathbf{f}_{P_i})} \mathbf{q} \\ \frac{1}{w_i(\mathbf{f}_{P_i})} \mathbf{q}^* & 1 \end{bmatrix} \succeq 0, \quad i = 1, \dots, p, \\ &\mathbf{q}_{T^c} = 0, \end{aligned} \quad (24)$$

where $\{\mathbf{G}_{0j}\}_{j=1}^p \in \mathbb{C}^{n^2 \times n^2}$ and $\{\{\mathbf{G}_{lj}\}_{l=1}^4\}_{j=1}^p \in \mathbb{C}^{(n-1)^2 \times (n-1)^2}$.

problem (24) generalizes the semidefinite programming (10) to the case that one has extra information about the locations of true signal frequencies. As our numerical experiments in the next section verifies, incorporation of prior information improves the frequency recovery performance.

III. NUMERICAL EXPERIMENT

In this section, we investigate performance of the proposed algorithm by numerical experiments. First, we generate a signal with chi-square distributed amplitude (i.e. $d \sim \delta_{0.5} + \chi^2(1)$) and uniform distributed phase $\phi \sim \mathcal{U}(0, 2\pi)$. We restrict our experiments to two cases. In the former, we assume that $p = 2, s = 4, n = 7$ and generate four frequency components uniformly in $[0, 0.2] \times [0, 0.2]$ and $[0.5, 0.7] \times [0.5, 0.7]$ without imposing any minimum separation condition. Measurements are uniform resampling of a set of 49 samples without replacement. We assume $\omega(f_{p1}) = \omega(f_{p2}) = 1$ and solve (24)

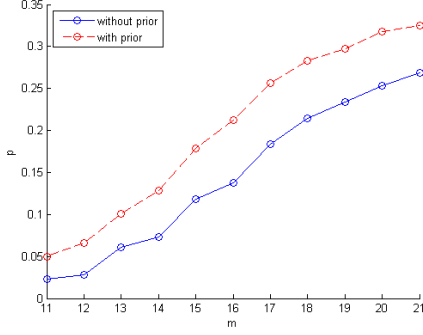


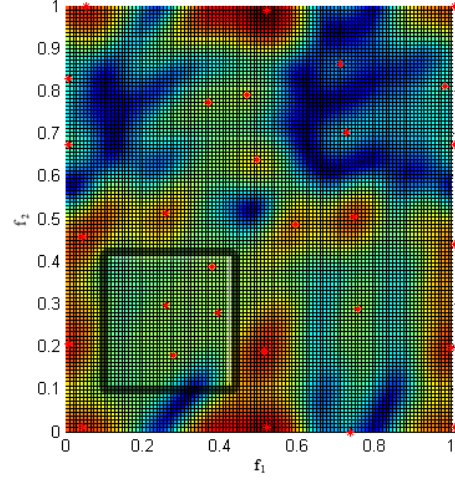
Fig. 2. Comparison of the proposed method with the case that there is no prior information. Experiments are done with 1000 trials, $n_1 = n_2 = 7, s = 4$ with the additional assumption that the true frequencies are uniformly distributed in $[0, 0.2] \times [0, 0.2]$ and $[0.5, 0.7] \times [0.5, 0.7]$ intervals.

using CVX [17] to recover the frequencies. It is shown in Fig. 2 that the recovery performance improves as we incorporate prior information.

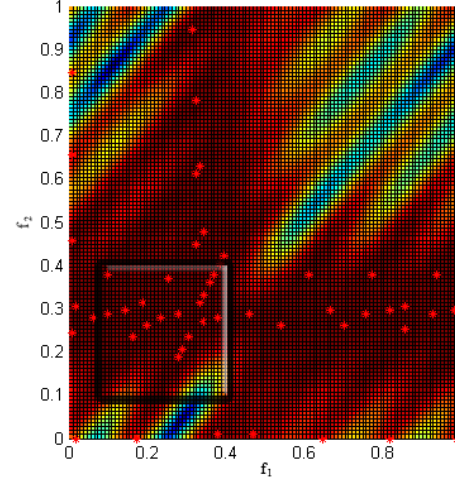
In the second simulation we assume $n = 7, p = 2$. Also, all of the measurements are available. We generate four frequencies uniformly in the domain $[0.1, 0.4] \times [0.1, 0.4]$ without any separation condition and assume there are prior knowledge in the domain and its complementary given by $P_F(\mathbf{f}_{p_1}) = 4.13$ and $P_F(\mathbf{f}_{p_2}) = 0.014$, respectively. Using this information, we obtain $w(\mathbf{f}_{p_1}) = 0.242$ and $w(\mathbf{f}_{p_2}) = 71.42$. Now, to recover the frequency components with and without prior knowledge, we solve (24) and (10), respectively. The dual polynomial of weighted atomic norm and regular atomic norm are plotted in Figs. 3(a) and 3(b), respectively. Fig. 3 demonstrates that even by observing all samples, regular atomic norm minimization can not recover the true frequencies properly while the proposed approach clearly identifies the frequency components.

REFERENCES

- [1] M. J. Rust, M. Bates, and X. Zhuang, “Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm),” *Nature methods*, vol. 3, no. 10, pp. 793–796, 2006.
- [2] M. Haardt, M. D. Zoltowski, C. P. Mathews, and J. Nosske, “2-d unitary esprit for efficient 2-d parameter estimation,” in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, vol. 3, pp. 2096–2099, IEEE, 1995.
- [3] Y. Hua, “Estimating two-dimensional frequencies by matrix enhancement and matrix pencil,” *IEEE Transactions on Signal Processing*, vol. 40, no. 9, pp. 2267–2280, 1992.
- [4] E. J. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Transactions on information theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [5] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [6] R. G. Baraniuk, “Compressive sensing [lecture notes],” *IEEE signal processing magazine*, vol. 24, no. 4, pp. 118–121, 2007.
- [7] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, “Sensitivity to basis mismatch in compressed sensing,” *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2182–2195, 2011.
- [8] A. Fannjiang and W. Liao, “Coherence pattern-guided compressive sensing with unresolved grids,” *SIAM Journal on Imaging Sciences*, vol. 5, no. 1, pp. 179–202, 2012.
- [9] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, “Compressed sensing off the grid,” *IEEE transactions on information theory*, vol. 59, no. 11, pp. 7465–7490, 2013.
- [10] Y. Chi and Y. Chen, “Compressive two-dimensional harmonic retrieval via atomic norm minimization,” *IEEE Transactions on Signal Processing*, vol. 63, no. 4, pp. 1030–1042, 2015.
- [11] Y. Chen and Y. Chi, “Robust spectral compressed sensing via structured matrix completion,” *Information Theory, IEEE Transactions on*, vol. 60, no. 10, pp. 6576–6601, 2014.
- [12] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, “The convex geometry of linear inverse problems,” *Foundations of Computational mathematics*, vol. 12, no. 6, pp. 805–849, 2012.
- [13] W. Xu, J.-F. Cai, K. V. Mishra, M. Cho, and A. Kruger, “Precise semidefinite programming formulation of atomic norm minimization for recovering d-dimensional (d 2) off-the-grid frequencies,” in *Information Theory and Applications Workshop (ITA), 2014*, pp. 1–4, IEEE, 2014.
- [14] M. A. Khajehnejad, W. Xu, A. S. Avestimehr, and B. Hassibi, “Weighted 1 minimization for sparse recovery with prior information,” in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, pp. 483–487, IEEE, 2009.
- [15] K. V. Mishra, M. Cho, A. Kruger, and W. Xu, “Spectral super-resolution with prior knowledge,” *IEEE Transactions on Signal Processing*, vol. 63, no. 20, pp. 5342–5357, 2015.
- [16] B. A. Dumitrescu, *Positive trigonometric polynomials and signal processing applications*. Springer Science & Business Media, 2007.



(a)



(b)

Fig. 3. The dual polynomial with $n_1 = n_2 = 7, s = 4, m = 49$. (a) with prior information about presence of frequencies in $[0.1, 0.4] \times [0.1, 0.4]$, (b) without prior information.

- [17] M. Grant, S. Boyd, and Y. Ye, “Cvx: Matlab software for disciplined convex programming,” 2008.